

# On the Sum of a Prime and a Square-free Number

ADRIAN W. DUDEK  
 Mathematical Sciences Institute  
 The Australian National University  
 adrian.dudek@anu.edu.au

## Abstract

We prove that every integer greater than two may be written as the sum of a prime and a square-free number.

## 1 Introduction

We say that a positive integer is square-free if it is not divisible by the square of any prime. It was first shown by Estermann [3] in 1931 that every sufficiently large positive integer  $n$  can be written as the sum of a prime and a square-free number. In particular, he proved that the number  $T(n)$  of such representations satisfies the asymptotic formula

$$T(n) \sim \frac{cn}{\log n} \prod_{p|n} \left(1 + \frac{1}{p^2 - p - 1}\right), \quad (1)$$

where

$$c = \prod_p \left(1 - \frac{1}{p(p-1)}\right) = 0.3739558 \dots \quad (2)$$

is a product over all prime numbers which is known as Artin's constant (see Wrench's computation [11] for more details). In 1935, Page [8] improved Estermann's result by giving a bound for the order of the error term in (1), using estimates for the error in the prime number theorem for arithmetic progressions. Mirsky [5] extended these results in 1949 to count representations of an integer as the sum of a prime and a  $k$ -free number, that is, a number which

is not divisible by the  $k$ -th power of any prime. More recently, in 2005, Languasco [4] treated the possibility of Siegel zeroes (see Davenport [2] for some discussion on this) with more caution so as to provide better bounds on the error.

The objective of the present paper is to prove the following theorem which completes the result of Estermann.

**Theorem 1.** *Every integer greater than two is the sum of a prime and a square-free number.*

We prove this theorem by working in the same manner as Estermann, though we employ explicit estimates on the error term for the prime number theorem in arithmetic progressions. Specifically, if we let

$$\theta(x; q, a) = \sum_{\substack{p \leq x \\ p \equiv a \pmod{q}}} \log p$$

where the sum is over primes  $p$ , we require estimates of the form

$$\left| \theta(x; q, a) - \frac{x}{\varphi(q)} \right| < \frac{\epsilon x}{\varphi(q)} \quad (3)$$

where  $\epsilon > 0$  is sufficiently small and  $x$  and  $q$  are suitably ranged. Good estimates of this type are available due to Ramaré and Rumely [10], but are only provided for finitely many values of  $q$ . It turns out, however, that this is sufficient, as the Brun–Titchmarsh theorem is enough for the remaining cases.

It should be noted that the best known asymptotic result in this area is that of Chen [1], who proved that every sufficiently large even number is the sum of a prime and another number which is the product of at most two primes. Chen’s result is, at present, the closest one seems to be able to get towards a proof of the Goldbach conjecture, which famously asserts that every even integer greater than two can be written as the sum of two primes. It would be interesting to see if one could modify the explicit proof of this paper to consider, instead of square-free numbers, numbers which are products of at most  $m$  primes (where  $m$  is fixed).

## 2 The Proof

### 2.1 The Setup

Let  $n \geq 10^{10}$  be a positive integer. We will prove Theorem 1 for this range before resorting to direct computation for the remaining cases. As convention

would have it, we let  $\mu : \mathbb{N} \rightarrow \{-1, 0, 1\}$  denote the Möbius function, where  $\mu(n)$  is zero if  $n$  is not square-free; otherwise  $\mu(n) = (-1)^{\omega(n)}$  where  $\omega(n)$  denotes the number of distinct prime factors of  $n$ . As such, we have that  $\mu(1) = 1$ . For a positive integer  $n$ , it can be shown that the sum

$$\mu_2(n) = \sum_{a^2|n} \mu(a)$$

is equal to 1 if  $n$  is square-free and zero otherwise. Thus, it follows that the expression

$$T(n) := \sum_{p \leq n} \mu_2(n - p)$$

counts the number of ways that  $n$  may be expressed as the sum of a prime and a square-free number. We will employ logarithmic weights so as to use the known prime number estimates with more ease, and so we define

$$R(n) := \sum_{p \leq n} \mu_2(n - p) \log p.$$

We note that  $n$  is the sum of a prime and a square-free number if and only if  $R(n) > 0$ . As such, the majority of this paper is dedicated to finding a lower bound for  $R(n)$ . The expression for  $R(n)$  can be rearranged so as to involve weighted sums over the prime numbers in arithmetic progressions:

$$\begin{aligned} R(n) &= \sum_{p \leq n} \log p \sum_{a^2|(n-p)} \mu(a) \\ &= \sum_{a \leq n^{1/2}} \mu(a) \sum_{\substack{p \leq n \\ a^2|(n-p)}} \log p \\ &= \sum_{a \leq n^{1/2}} \mu(a) \theta(n; a^2, n). \end{aligned}$$

We will split the range of this sum into three parts, for we shall use a different technique to bound each of them. Note first that if  $(a, n) > 1$ , then we have trivially that  $\theta(n; a^2, n) \leq \log n$ . Thus, we may write

$$R(n) > \Sigma_1 + \Sigma_2 + \Sigma_3 - n^{1/2} \log n \tag{4}$$

where

$$\begin{aligned}\Sigma_1 &= \sum_{\substack{a \leq 13 \\ (a,n)=1}} \mu(a)\theta(n; a^2, n), \\ \Sigma_2 &= \sum_{\substack{13 < a \leq n^A \\ (a,n)=1}} \mu(a)\theta(n; a^2, n), \\ \Sigma_3 &= \sum_{\substack{n^A < a \leq n^{1/2} \\ (a,n)=1}} \mu(a)\theta(n; a^2, n),\end{aligned}$$

and  $A \in (0, 1/2)$  is to be chosen later to optimise our result. We will use the estimates of Ramaré and Rumely [10] to bound  $\Sigma_1$ ; this is the reason for the specific range of  $a$  in this sum. We will then use the Brun–Titchmarsh theorem to bound  $\Sigma_2$ . Finally,  $\Sigma_3$  will be bounded using trivial estimates.

## 2.2 Arithmetic estimates

We first consider the sum

$$\Sigma_1 = \sum_{\substack{a \leq 13 \\ (a,n)=1}} \mu(a)\theta(n; a^2, n).$$

Theorem 1 of Ramaré and Rumely [10] provides estimates of the form

$$\left| \theta(n; a^2, n) - \frac{n}{\varphi(a^2)} \right| < \frac{\epsilon_a n}{\varphi(a^2)}. \quad (5)$$

In particular, by looking through the square moduli in Table 1 of their paper, we have values of  $\epsilon_a$  for all  $1 \leq a \leq 13$  which are valid for all  $n \geq 10^{10}$ . We therefore have trivially that

$$\begin{aligned}\Sigma_1 &> n \sum_{\substack{a \leq 13 \\ (a,n)=1}} \left( \frac{\mu(a)}{\varphi(a^2)} - \epsilon_a \frac{\mu^2(a)}{\varphi(a^2)} \right) \\ &> n \left( \sum_{(a,n)=1} \frac{\mu(a)}{\varphi(a^2)} - \sum_{\substack{a > 13 \\ (a,n)=1}} \frac{\mu(a)}{\varphi(a^2)} - \sum_{a \leq 13} \frac{\epsilon_a \mu^2(a)}{\varphi(a^2)} \right).\end{aligned} \quad (6)$$

We wish to estimate the three sums in the above parentheses. We denote the leftmost sum by  $S_n$ , and note that we can bound it below by Artin’s constant

(2) *viz.*

$$\begin{aligned}
S_n &:= \sum_{(a,n)=1} \frac{\mu(a)}{\varphi(a^2)} = \prod_{p \nmid n} \left(1 - \frac{1}{\varphi(p^2)}\right) \\
&= \prod_{p \nmid n} \left(1 - \frac{1}{p(p-1)}\right) \\
&\geq \prod_p \left(1 - \frac{1}{p(p-1)}\right) = c.
\end{aligned}$$

Wrench [11] has computed this constant to high accuracy; it will suffice for the purpose of Theorem 1 to note that  $S_n \geq c > 0.373$ .

We will, for the moment, neglect the middle sum in (6), for it shall be considered jointly with a term in the estimation of  $\Sigma_2$ . Thus, in our estimation of  $\Sigma_1$ , it remains to manually compute the upper bound for the rightmost sum. This is a straightforward task which is done in reference to Table 1 of Ramare and Rumely's paper [10]. We get that

$$\sum_{a \leq 13} \frac{\epsilon_a \mu^2(a)}{\varphi(a^2)} < 0.005.$$

We now bring  $\Sigma_2$  into the fray; the explicit Brun–Titchmarsh theorem (see Montgomery and Vaughan [6]) provides the bound

$$\theta(n; a^2, n) < 2 \left( \frac{\log n}{\log(n/a^2)} \right) \frac{n}{\varphi(a^2)}.$$

Clearly, in the range  $13 < a \leq n^A$  we may bound

$$\frac{\log n}{\log(n/a^2)} \leq \frac{1}{1-2A}$$

and so we have the estimate that

$$\theta(n; a^2, n) = \frac{n}{\varphi(a^2)} + \epsilon \left( \frac{1+2A}{1-2A} \right) \frac{n}{\varphi(a^2)}$$

where  $|\epsilon| < 1$ . We may then bound  $\Sigma_2$  from below by

$$\Sigma_2 > n \sum_{\substack{13 < a \leq n^A \\ (a,n)=1}} \frac{\mu(a)}{\varphi(a^2)} - n \left( \frac{1+2A}{1-2A} \right) \sum_{\substack{13 < a \leq n^A \\ (a,n)=1}} \frac{\mu^2(a)}{\varphi(a^2)}.$$

We can then add this to our estimate for  $\Sigma_1$  to get

$$\begin{aligned}
\Sigma_1 + \Sigma_2 &> n \left( S_n - 0.005 - \sum_{\substack{a > n^A \\ (a,n)=1}} \frac{\mu(a)}{\varphi(a^2)} - \left( \frac{1+2A}{1-2A} \right) \sum_{\substack{13 < a \leq n^A \\ (a,n)=1}} \frac{\mu^2(a)}{\varphi(a^2)} \right) \\
&> n \left( S_n - 0.005 - \left( \frac{1+2A}{1-2A} \right) \sum_{a > 13} \frac{\mu^2(a)}{\varphi(a^2)} \right). \tag{7}
\end{aligned}$$

We can estimate the sum in the above inequality by writing it as follows:

$$\sum_{a > 13} \frac{\mu^2(a)}{\varphi(a^2)} = \sum_{a=1}^{\infty} \frac{\mu^2(a)}{\varphi(a^2)} - \sum_{a \leq 13} \frac{\mu^2(a)}{\varphi(a^2)}.$$

The infinite sum is less than 1.95 (see Ramaré [9] for example), and the finite sum can be computed by hand to see that the sum in (7) is bounded above by 0.086. Thus

$$\Sigma_1 + \Sigma_2 > n \left( S_n - 0.005 - \left( \frac{1+2A}{1-2A} \right) (0.086) \right). \tag{8}$$

For  $\Sigma_3$ , we have trivially that

$$\begin{aligned}
|\Sigma_3| &< \sum_{n^A < a \leq n^{1/2}} \theta(n; a^2, n) \\
&< \sum_{n^A < a \leq n^{1/2}} \left( 1 + \frac{n}{a^2} \right) \log n \\
&< n^{1/2} \log n + n \log n \sum_{n^A < a \leq n^{1/2}} \frac{1}{a^2} \\
&< n^{1/2} \log n + n^{1-2A} \log n + n \log n \int_{n^A}^{n^{1/2}} t^{-2} dt \\
&= n^{1-2A} \log n + n^{1-A} \log n. \tag{9}
\end{aligned}$$

### 2.3 A Lower Bound for $R(n)$

We can now provide an explicit lower bound for

$$R(n) \geq \Sigma_1 + \Sigma_2 + \Sigma_3 - n^{1/2} \log n.$$

We combine our explicit estimates (8) and (9) and divide through by  $n$  to get that

$$\begin{aligned} \frac{R(n)}{n} &> S_n - 0.005 - \left( \frac{1+2A}{1-2A} \right) (0.086) \\ &- n^{-1/2} \log n - n^{-2A} \log n - n^{-A} \log n. \end{aligned}$$

For sufficiently small  $A$  and large  $n$ , the right hand side will be positive. For any  $n$ , we have  $S_n > 0.373$ ; it is a simple matter to choose  $A = 1/4$  and verify that the right hand side is positive for all  $n \geq 10^{10}$ . That is, Theorem 1 is true for all integers  $n \geq 10^{10}$ .

It so remains to prove this result for all integers in the range  $3 \leq n < 10^{10}$ . If  $n$  is even, we have the numerical verification by Oliviera e Silva, Herzog and Pardi [7] that all even integers up to  $4 \cdot 10^{18}$  can be written as the sum of two primes. Thus, every even integer greater than two may be written as the sum of a prime and a square-free number.

Thus, we need to check that every odd integer  $3 \leq n \leq 10^{10}$  can be written as the sum of a prime and a square-free number. For each  $n$ , we subtract a prime number  $p$  and check that the result is squarefree. This is a straightforward computation; we ran this on MATHEMATICA and it took just under 3 days on a 2.6GHz laptop. The computation was eased somewhat by subtracting primes which were close in size to  $n$ .

## Acknowledgements

The author wishes to thank Dr Timothy Trudgian for many helpful conversations regarding the present paper.

## References

- [1] J.-R Chen. On the representation of a larger even integer as the sum of a prime and the product of at most two primes. *Sci. Sinica*, 16:157–176, 1975.
- [2] H. Davenport. *Multiplicative number theory*. Springer-Verlag, New York, third edition, 2000.
- [3] T. Estermann. On the representations of a number as the sum of a prime and a quadratfrei number. *J. London Math. Soc.*, S1-6(3):219, 1931.
- [4] A. Languasco. On the sum of a prime and a  $k$ -free number. *Funct. Approx. Comment. Math.*, 34:19–26, 2006.
- [5] L. Mirsky. The number of representations of an integer as the sum of a prime and a  $k$ -free integer. *Amer. Math. Monthly*, pages 17–19, 1949.
- [6] H. L. Montgomery and R. C. Vaughan. The large sieve. *Mathematika*, 20(02):119–134, 1973.
- [7] S. Oliveira e Silva, T. Herzog and S. Pardi. Empirical verification of the even Goldbach conjecture and computation of prime gaps up to  $4 \cdot 10^{18}$ . *Math. Comp.*, 83(288):2033–2060, 2014.
- [8] A. Page. On the number of primes in an arithmetic progression. *Proc. Lond. Math. Soc.*, 2(1):116–141, 1935.
- [9] O. Ramaré. On Šnirel’man’s constant. *Ann. Sc. Norm. Super. Pisa Cl. Sci.*, 22(4):645–706, 1995.
- [10] O. Ramaré and R. Rumely. Primes in arithmetic progressions. *Math. Comp.*, 65(213):397–425, 1996.
- [11] J. W. Wrench. Evaluation of Artin’s constant and the twin-prime constant. *Math. Comp.*, pages 396–398, 1961.